AGENT-BASED Modeling and Simulation for EXASCALE Computing

Researchers at Argonne National Laboratory have been successfully using a new modeling paradigm—agent-based modeling and simulation (ABMS)—to address challenges and gain valuable insights in such key areas as energy, biology, economics, and social sciences. To maximize potential, they are developing a next-generation ABMS system that can be extended to exascale computing environments to achieve breakthrough results in science, engineering, and policy analysis.

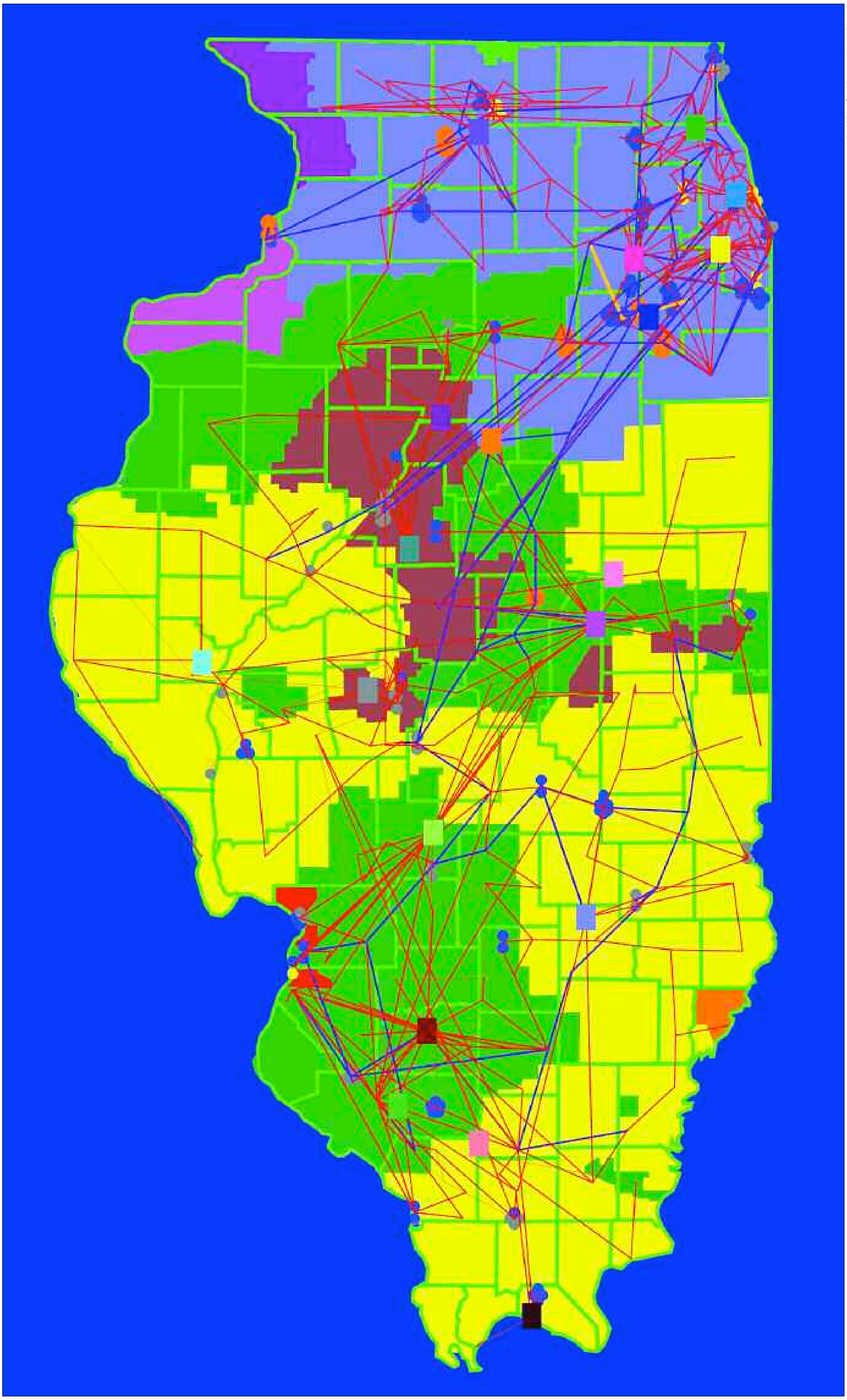
Argonne researchers have developed and used large-scale agent-based models to provide important information to policymakers that would not be available using other modeling approaches.

Argonne National Laboratory (ANL) is a leader in agent-based modeling and simulation (ABMS). ABMS is a new modeling paradigm that is having far-reaching effects on the way that researchers across disciplines use electronic laboratories to conduct their research. By modeling systems from the ground up, researchers are exploring how system behaviors emerge from the behaviors of large numbers of interacting individuals, or agents. ABMS also serves as an experimental technique, a framework for developing electronic laboratories in which the most detailed assumptions about individual agents, their behaviors, and interactions can be varied and explored *in silico*.

Computational advances have opened the way for a growing number of agent-based applications across many fields. These applications range from modeling adaptive behaviors and the emergence of new entities in the biological sciences (sidebar "ABMS Benefits Biological Sciences," p37) to modeling agent behavior in the stock market and supply chains to understanding consumer purchasing (sidebar "Agent-Based Modeling Applications," p40).

ABMS provides new ways for businesses and government to use computers to support decision making and to analyze policies. For social systems that are composed of agents who learn and adapt their behavior based on their individual experiences, ABMS explores how decisions and policies may affect groups and individuals before the decisions are made or the policies are implemented.

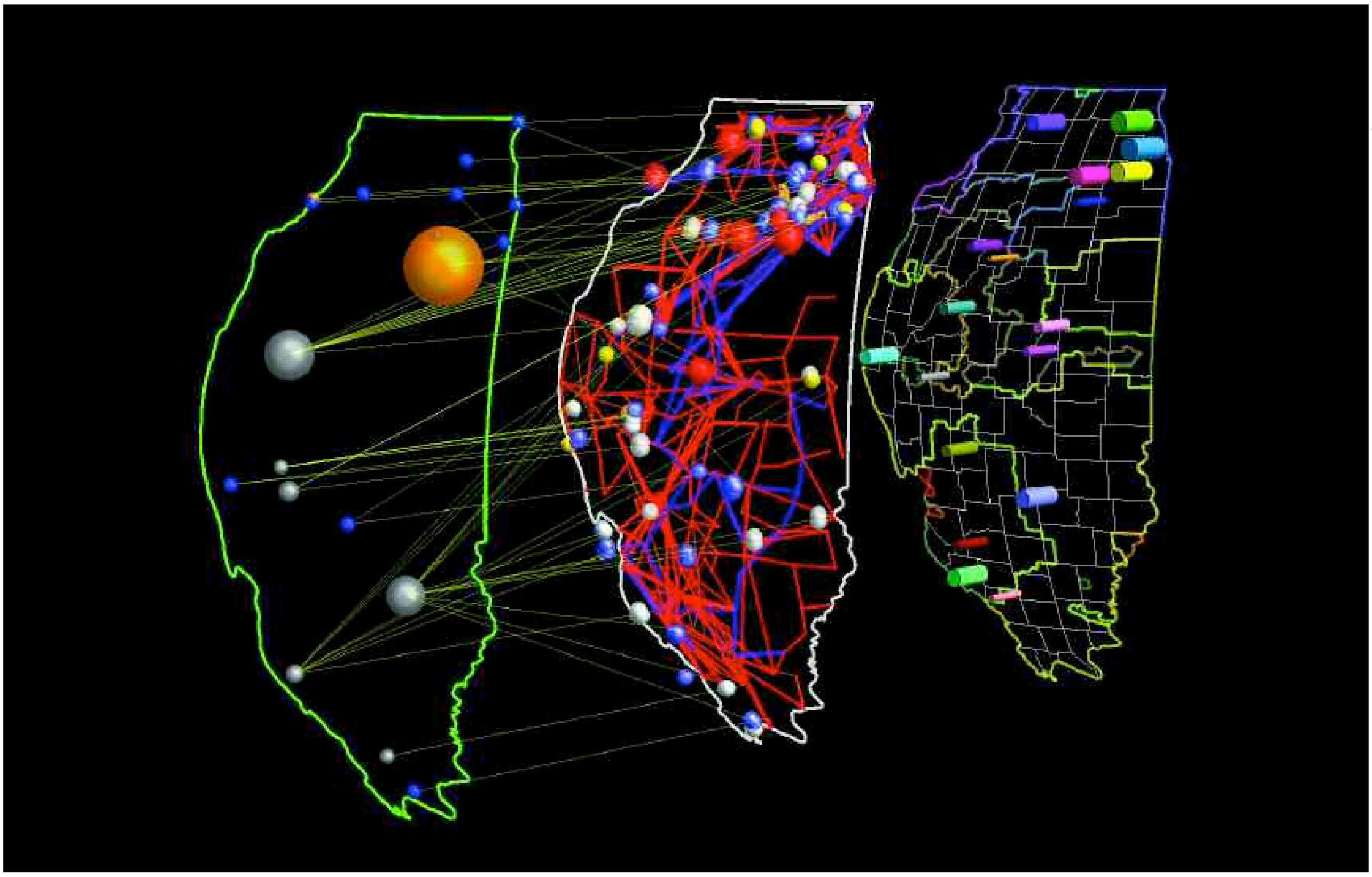
Argonne researchers have developed and used large-scale agent-based models to provide important information to policymakers that would not be available using other modeling approaches. One outstanding example—Electricity Markets Complex Adaptive Systems (EMCAS)—was used to model the Illinois electric power industry under deregulation conditions in an effort to anticipate the likely effects of deregulation on electricity prices and reliability. In this model, a



C. Macal and M. North, Argonne

Electricity Markets Complex Adaptive Systems (EMCAS) was used to model the Illinois electric power industry under deregulation conditions in an effort to anticipate the likely effects of deregulation on electricity prices and reliability.

Figure 1. *Illinois electric power transmission grid and service areas.*



C. Macal and M. North, Argonne

Figure 2. *Illinois electric power market elements, showing generating companies and ownership relationships (left), electric generators and transmission network (center), and service area loads (right).*

The growing roster of successful ABMS applications demonstrates the enormous potential of this technique.

detailed representation of the electric power grid was combined with a representation of electric power market agents acting in a deregulated market environment. Generating company agents adapted in response to changing market conditions in the simulation as would be expected of agents in the real world (figure 1, p35; figure 2)

Expanding ABMS Use

Agent-based modeling is becoming widespread for a number of reasons. First, the systems that need to be analyzed and modeled are becoming more complex in their structure and interdependencies. Traditional modeling tools are no longer as applicable as they once were. Second, some systems have always been too complex to adequately model. For instance, when modeling economic markets, researchers often relied on the notions of perfect information, homogeneous agents, and long-run equilibrium because these assumptions made the problems analytically and computationally solvable. Using ABMS, they are starting to relax some of these assumptions and to understand the effect of these simplifying assumptions on models of economic systems. Third, data are being collected and organized into databases at finer levels of granularity. These micro-data can now support individual-based simulations. Fourth, and most important, computational power is advancing rapidly. Researchers can now run large-scale micro-simulation models that would not have been computationally executable just a couple of years ago.

Remarks Dr. Charles Macal, director of Argonne's Center for Complex Adaptive Agent Systems Simulation, "Agent-based models often produce what have been referred to as emergent phenomena—nonlinear effects that cannot be anticipated without resorting to computation and that are the direct result of many agents repeatedly interacting, using only local information."

Exploring ABMS on the Exascale

The growing roster of successful ABMS applications demonstrates the enormous potential of this technique (sidebar "Agent-based Modeling Applications," p40). However, realizing its full potential to achieve breakthrough results in science, engineering, and policy analysis requires far greater computing capability than is possible through current tools and approaches. While many agent-based software packages and toolkits are available (sidebar "Agent-Based Software

ABMS studies typically require the execution of many model runs.

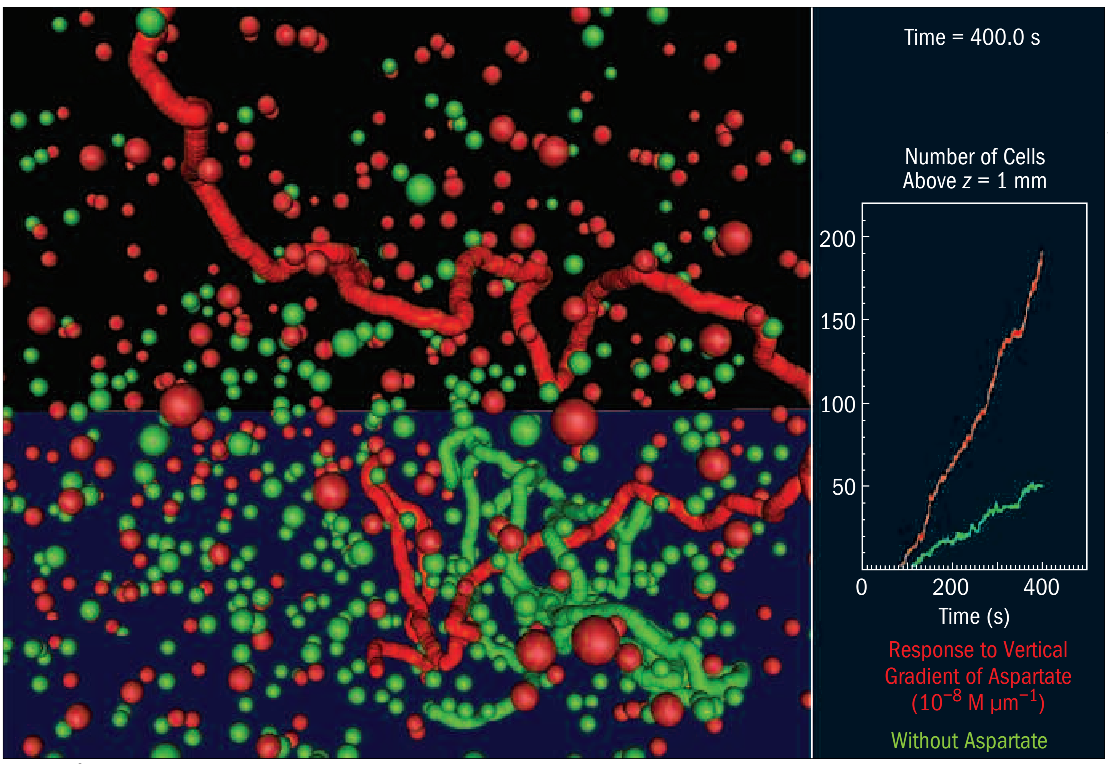
ABMS Benefits Biological Sciences

In the biological sciences, Argonne researchers are using ABMS to model cellular behavior. Using Argonne's high-performance Jazz cluster, researchers developed AgentCell, an agent-based simulator for modeling the chemotactic processes involved in the motile behavior of the *Escherichia coli* bacteria. AgentCell is an open-source tool, based on the Repast Simphony toolkit developed by Argonne and University of Chicago researchers. The research examined how the range of natural cell diversity is responsible for the full range of cell behaviors (figure 3).

The researchers used Jazz to run numerous simulations of the *E. coli* chemotaxis network for many cells and stochastic variations. They modeled agents at the molecular level as well as at the whole cell level. Argonne is now investigating the efficient implementation of agent-to-agent interactions that minimize resource contention through its exascale computing research effort.

Large-scale simulations of digital bacteria run on the Jazz computing cluster suggested that the chemotaxis network is tuned to simultaneously optimize the random spread of cells in the absence of nutrients and the cellular response to gradients of nutrients. Without the computing time on Jazz, this work would have been very difficult to complete. These computations have paved the way for closer collaborations between analytical treatments of, computational modeling of, and wet lab experimentation with *E. coli* behavior, which potentially has broad implications for many biological systems.

AgentCell
http://www.agentcell.org



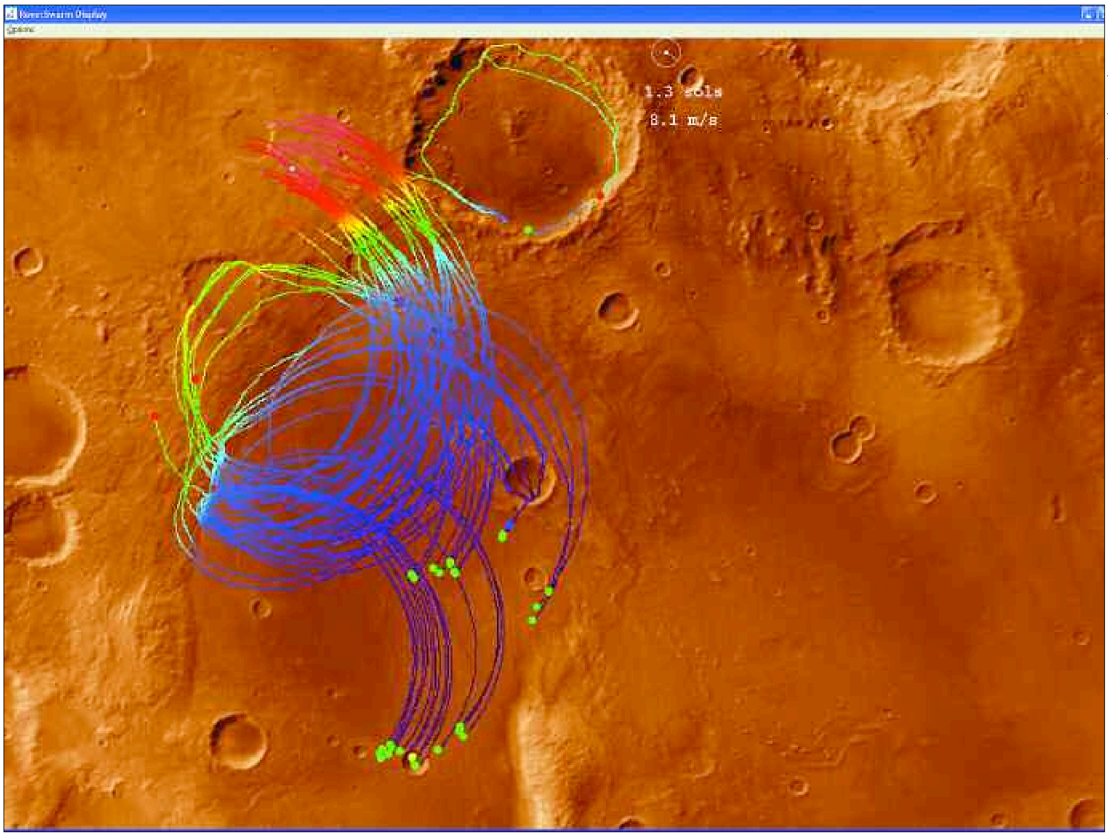
C. Macal and M. North, Argonne

Figure 3. *Cell trajectories in the AgentCell simulation.*

Toolkits," p41), these implementations are not aimed at extremely high-performance computing platforms. Recent efforts to increase the computational capacity of ABMS toolkits have focused on either simple parameter sweeps or the use of standardized, but performance-limited technologies.

ABMS studies typically require the execution of many model runs. Large numbers of model runs are needed to account for stochastic or random variation in model output results as well as to explore the possible range of outcomes that can be produced by the model in question. These ensembles of runs are often organized into parameter sweeps in which each model run has slight variations in inputs or random stream seeds. According to Dr. Ian Foster, director of the University of Chicago/Argonne Computation Institute, parameter sweeps are often described as "embarrassingly parallel" because of the way

NASA used Repast for an agent-based simulation on autonomous robots roaming the Martian surface.



K. Severance, NASA; K. Zick, University of Michigan.

Figure 4. *Swarms of autonomous rovers establish communication links and coordinate movements over the surface of a virtual Mars in this Repast simulation of self-organizing agent behavior.*

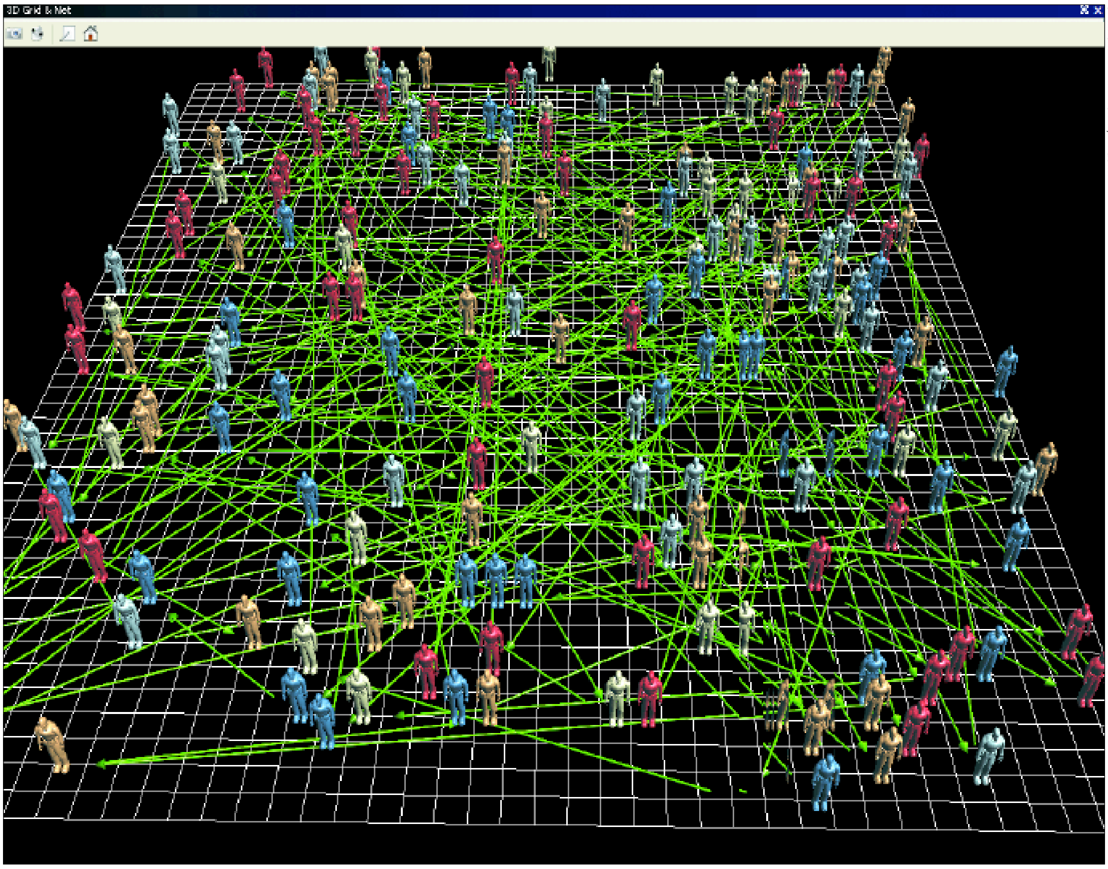
that the independence of the individual model runs allows them to be easily distributed over a large number of processors on high-performance computers. Quite a few computing groups have developed tools for performing such parameter sweeps on high-performance computing systems. However, many of the most interesting agent-based modeling systems require extremely large individual model runs rather than large numbers of small model runs. The parameter sweep approach does not offer a solution to this problem. Other widely used standards, such as the High Level Architecture and various kinds of web services, are also not designed to work efficiently on extremely large computing platforms.

To address the pressing need for extremely large-scale agent-based models, Argonne researchers are exploring a next-generation ABMS system that can be extended to exascale computing environments. Explains Dr. Michael North, deputy director of Argonne's Center for Complex Adaptive Agent Systems Simulation, "Before we can efficiently use high-performance computing, we need to do research to understand how to best use the architectures. Switching our model to run on 10,000 nodes instead of 10 is not just a linearly scalable process. Moving from a few interactions to a million interactions is not just moving to greater numbers. Entire interactions change. Also, we need to consider implementation issues such as partitioning of the name space, memory space, communication, and the effect of noise and jitter on these large platforms. Ultimately, we need to make changes to our model to deal with such issues."

The researchers plan to do initial test runs on the smaller of the two Blue Gene/P systems at the Argonne Leadership Computing Facility (ALCF; "Opening New Possibilities in Science and Engineering," *SciDAC Review*, Spring 2008, p42). The exascale ABMS system being developed will leverage existing world-class capabilities such as Argonne's Repast Simphony ABMS toolkit and distributed computing expertise to target multi-core chips. The project aims to scale to 106 or 107 threads. A thread is an independent instruction execution flow that can share resources, such as memory, with other threads in the same parent process.

Pilot Applications for Exascale ABMS

Three pilot application areas—microbial biodiversity, cybersecurity, and the social aspects of climate change—are providing a portfolio of requirements for exascale ABMS. Each application area suggests a unique series of experiments that cumulatively cover the range of functionality



C. Macal and M. North, Argonne

Figure 5. *A Repast dynamic social network simulation.*

required for an agent-based architecture. Depending on the findings of the experiments, a variety of possible designs and implementation paths will be considered. Recent Argonne work on the Repast Simphony ABMS toolkit offers one possible path among several that will be experimentally tested. This toolkit has been used successfully for a wide range of applications. For example, the National Aeronautics and Space Administration (NASA) used Repast for an agent-based simulation on autonomous robots roaming the Martian surface (figure 4) and a dynamic social network simulation from Repast is shown in figure 5.

Enabling large-scale simulations to address the complexity of real microbial environments is a major focus of this research. "Scalable ABMS simulations show great promise for understanding the detailed dynamics of large, mixed microbial communities, with a wide range of applications in ecology, health sciences, and industry," says Rick Stevens, associate laboratory director for Computing, Environment, and Life Science at Argonne. The rapid accumulation of environmental molecular data is uncovering vast diversity, abundant uncultivated microbial groups, and novel microbial functions. This accumulation of data requires the application of theory and simulation to provide organization, structure, insight, and, ultimately, predictive power that is of practical value. Argonne researchers are using resource-ratio theory in microbial ecosystems in devising requirements for the practical application of agent-based modeling to the exascale system under development.

Cybersecurity also offers considerable opportunities to apply exascale ABMS. There is a need to develop advanced methods to explore future cybersecurity scenarios that involve a variety of defensive strategies. Simulation techniques are one approach that may yield insight to vulnerabilities and security dynamics. The scope of simulation includes the systems—such as computers, networks, routers, filters, and monitors— users of these systems, administrators, configuration processes, cybersecurity policies, and external threat agents.

The information technology (IT) infrastructures are large-scale systems with 106 to 1,010 IT entities consisting of a few hundred types. Each entity can have a very large number of possible internal states and many thousands of users and administrators. These systems are geographically dispersed and connected by complex networking with constantly evolving topologies. Agent-based modeling is a natural methodology for simulating such infrastructures since both the

Ultimately, the goal is to develop an open-source agent-based modeling toolkit that is highly scalable.

Agent-based Modeling Applications

ABMS is being actively applied in many practical areas. The applications range across a continuum, from elegant, minimalist academic models to large-scale decision support systems.

Minimalist models are based on a set of idealized assumptions, designed to capture only the most salient features of a system. These agent-based models are exploratory electronic laboratories in which a wide range of assumptions can be varied over a large number of simulations.

Decision support models tend to serve large-scale applications and are designed to answer real-world policy questions. These models normally include real data and have passed appropriate validation tests to establish credibility.

Business and Organizations
- Consumer markets
- Supply chains
- Insurance
- Manufacturing

Economics
- Artificial financial markets
- Trade networks

Infrastructure
- Transportation
- Electric power markets
- Hydrogen economy

Crowds
- Human movement
- Evacuation modeling

Society and Culture
- Ancient civilizations
- Civil disobedience

Terrorism
- Social determinants
- Organizational networks

Military
- Command and control
- Combat
- Logistics

Biology
- Ecological systems
- Animal behavior
- Cell behavior
- Subcellular molecular behavior

technical dimensions (for example, the server protocols) and the human dimensions (such as using "social engineering" to gain user passwords) of these systems can be simultaneously modeled. Argonne researchers are developing prototype models of typical IT infrastructures and the associated mechanisms to explore current and alternative administrative policies. The linking of human behavior with network structure offers the possibility of breakthroughs in the fundamental understanding of cybersecurity.

Improving scientific understanding of climate change requires researchers to consider the physical, economic, and social determinants of climate change. Modeling all three aspects is essential for long-term climate forecasting since economic and social decision-making has a potentially important effect on physical climate factors (such as individual fuel choices and how use of energy technologies may affect atmospheric carbon dioxide concentrations). Similarly, physical climate factors may influence decision making (for example, rising atmospheric carbon dioxide concentrations may encourage people to reduce carbon dioxide emissions). Much progress has been made in modeling the physical aspects of climate change such as atmospheric flows, oceanic circulation, and albedo effects. Recent research in agent-based economic and social modeling has opened the door to modeling the feedback loops inherent in the social aspects of climate change. The resulting scenarios will require modeling up to 10^{10} human agents. These agents would each have a large number of complex internal states (10^3 or more) and be drawn from a wide range of behavioral types. Figure 6 illustrates the kind of scale-up being sought in the context of the example applications.

Ultimately, the goal is to develop an open-source agent-based modeling toolkit that is highly scalable. The agent-based modeling toolkit being developed for the exascale ABMS system has four major architectural components: a time scheduler, a storage system for agent-endogenous data, a storage system for agent topologies, and a data logging system. The storage systems for agent-endogenous data and agent topologies are generally synchronized to form a single data storage framework.

Time schedulers coordinate and synchronize the flow of agent activities as events unfold in the simulation. Storage systems for agent-endogenous data hold the internal state information for each agent (for example, each agent's scalar attributes). This component is also responsible for saving the agent's internal state information across executions of the model (such as storing input data or allowing runs to be replicated). Storage and access systems for agent topologies hold both the spatial and aspatial relationships between agents and manage the communications between agents that occur across these links. These systems are also responsible for saving the relationship information across separate executions of the model.

Argonne researchers plan to complete the backbone of the exascale ABMS system within the next year.

Agent-Based Software Toolkits

Agent modeling can be accomplished by using a desktop computer, computing clusters, or, potentially, grids. Typically, desktop agent-based models do not scale well to what is required for extremely large applications. Argonne's exascale computing research is working to break down the barriers.

The object-oriented modeling paradigm is the basis for agent modeling, since an agent can be considered a self-directed object with the capability of autonomously choosing actions based on the agent's situation. It is natural to use classes and methods to represent agents and agent behaviors.

Swarm was the first software tool created for agent-based modeling and simulation. Originally developed at the Santa Fe Institute in 1994, Swarm was specifically designed for Artificial Life applications. Since that time, there has been a steady advancement of the software toolkits and development environments that have superseded Swarm.

One example is Repast Simphony, a free and open-source toolkit developed by Argonne experts. The latest version has tools for visual model development, visual model execution, automated database connectivity, automated output logging, and results visualization.

Repast Simphony
http://repast.sourceforge.net/

Number of Processors	Bacteria Agents		Network Agents		Human Agents	
	Basic	Detailed	Basic	Detailed	Basic	Detailed
10^0	10^5	10^5	10^3	10^1	10^4	10^0
10^6	10^{11}	10^{11}	10^9	10^7	10^{10}	10^6
10^7	10^{12}	10^{12}	10^{10}	10^8	10^{11}	10^7
10^8	10^{13}	10^{13}	10^{11}	10^9	10^{12}	10^8

Figure 6. *Targeted scale-up in terms of number and complexity of agents.*

Data logging systems record the activities within and the results of simulations (for example, who talked to whom, and when, in a social network). The information stored by these systems is used for simulation analysis and visualization.

Argonne researchers plan to complete the backbone of the exascale ABMS system within the next year. They will then work to extend the system, based on a focused series of experiments in each of the pilot application areas, as well as furthering the development of the models for each domain. The system will enable the researchers to work with microbiologists, social scientists, and cybersecurity experts in advancing breakthrough science in these areas. In broader terms, the architecture and knowledge provided by this research will be applicable to many other areas as well. ●

Contributors: Dr. Charles M. Macal, director for Argonne's Center for Complex Adaptive Agent Systems Simulation; Dr. Michael J. North, deputy director for Argonne's Center for Complex Adaptive Agent Systems Simulation; Dr. Gail Pieper, senior coordinator of writing and editing for Argonne's Mathematics and Computer Science Division; and Cheryl Drugan, coordinating writer/editor for the Argonne Leadership Computing Facility

Further Reading

Complex Adaptive Systems
http://www.dis.anl.gov/exp/cas/index.html

Argonne Leadership Computing Facility
http://www.alcf.anl.gov

M. J. North and C. M. Macal. 2007. *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation.* Oxford University Press, Inc., New York, NY.

C. M. Macal and M. J. North. 2007. Tutorial on Agent-Based Modeling and Simulation: Desktop ABMS. *Proceedings of the 2007 Winter Simulation Conference.* S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew and R. R. Barton (eds.), pp. 95-106, Washington, DC, December 9-12, 2007. Paper available at: http://www.informs-sim.org/wsc07papers/011.pdf